

Weekly Report

2017.05.02-2017.05.07

1. This Week

Wavelines

- 1.learn about the coding details of the wavelines system from its source code
- 2.read the book “statistical control theory” and learn about the statical foundation of the EWMA control chart method
- 3.write the system using instructions
- 4.read several paper about anomaly detection and design.

Security Project

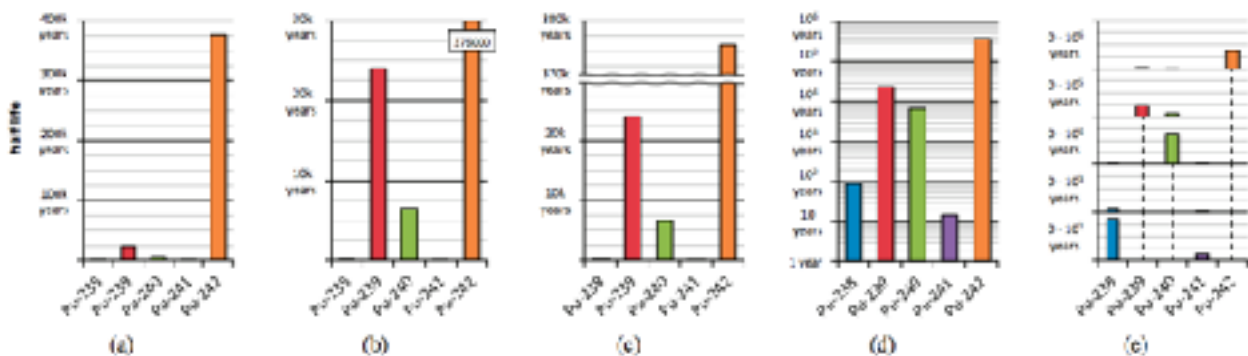
- 1.A set of new requirements has been asked for after testing by doc TIAN. I'll do my part after Huihua have made her arrangement.

Others

- 1.prepare and report the paper ‘Visplause’ during the group meeting
- 2.write the group blog about the reported paper

paper reading

1.Scale-Stack Bar Charts



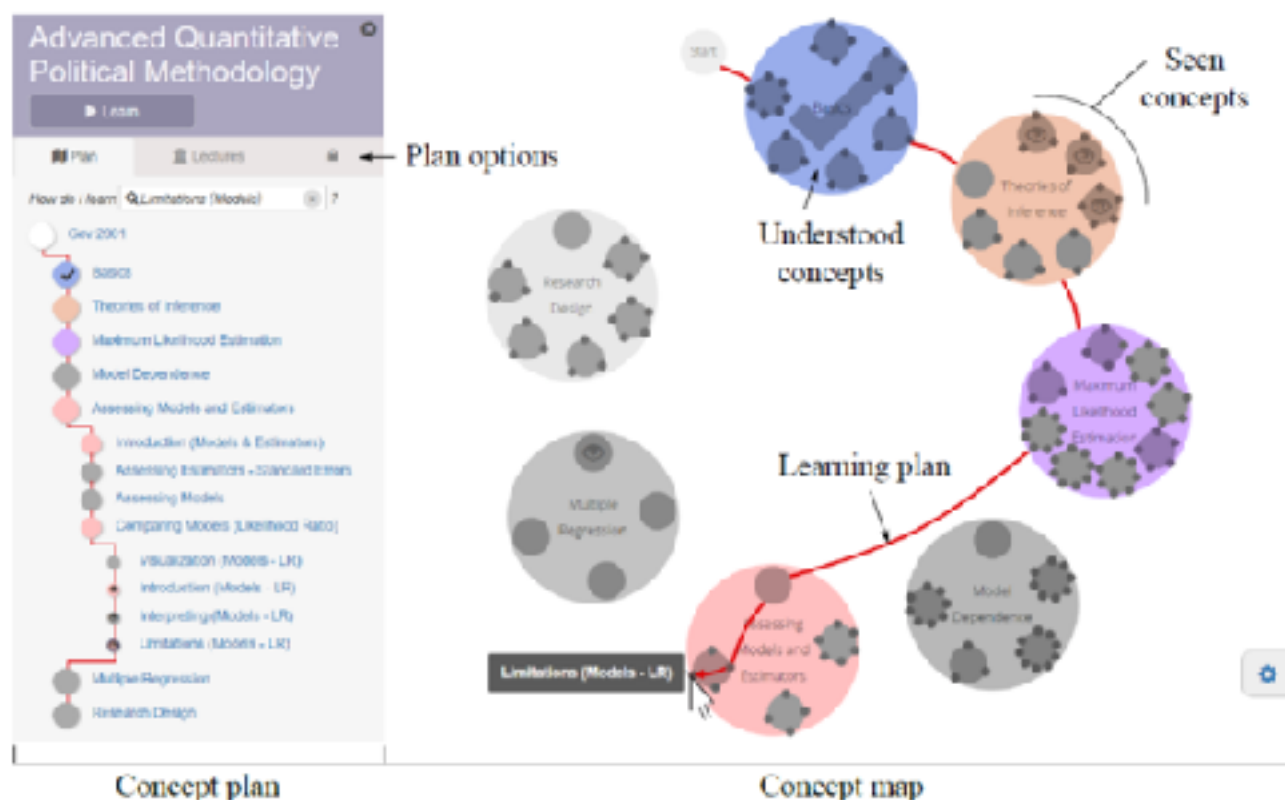
This paper presents a novel method to visualize large value ranges with bar charts: scale-stack bar charts (figure (e)). Since linear bar chart and advanced non-linear bar chart may cause the readability problem and make it difficult to compare the values of two categories that differ from each other a lot. Scale-stack bar chart split the area into 5 parts and each part have the range started from 0 to 10^n . This design works well when quantitatively analyzing data with large value ranges and suits for comparison tasks.

2.Anomaly Detection: A Survey

This survey classifies the recent anomaly problems and applications and then introduce the basic anomaly detection techniques: classification-based, nearest neighbor-based, clustering-based, statistical techniques, information theory, and spectral theory. Among all these methods, we are going to focusing on statistical methods. The challenges include:

- defining a normal region
- normal behavior keeps evolving
- rely on the assumption that the data is generated from a particular distribution
- choosing the best statistic
- implement Histogram-based techniques that suits for multivariate data

3.booc.io: An Education System with Hierarchical Concept Maps and Dynamic Non-linear Learning Plans



This paper allows linear and non-linear presentation and navigation of educational concepts and material. It provides a concept plan(left) and a concept map(right) for each course adapted to booc.io. The concept map focuses on course concepts learning which uses a circular layout to arrange concepts and their sub-concepts in a clockwise order. The concept plan acts as a search panel, a control panel, and a hierarchical list representation of the course. This paper is enlightening for hierarchical data visualization.

4.Profiler: Integrated Statistical Analysis and Visualization for Data Quality Assessment.

This paper includes modular components for custom data types, anomaly detection routines and summary visualization. It first analyze the input table is to infer types for each column. And then it use type information to generate features prior to running anomaly detection routines. Finally, the results of anomaly detection and mutual information analysis are used to perform view recommendation. It is attractive to first see recommendation mechanism in anomaly detection works.

Schema Browser

- Creative Type
- Distributor
- IMDb Rating
- IMDb Votes
- MPAA Rating
- Major Genre
- Production Budget

Related Views: Anomalies

Anomaly Browser

Missing (6)

- MPAA Rating
 - Creative Type
 - Source
 - Major Genre
 - Distributor
 - Release Location
- ### Error (2)
- Distributor (Levenshtein)
 - Source (Levenshtein)

Extreme (7)

Inconsistent (3)

